Discrete Applied Mathematics 10 (1985) 211-225
North-Holland

ON THE NUMBER OF CYCLES POSSIBLE IN DIGRAPHS WITH LARGE GIRTH*

Eric W. ALLENDER
School of Information and Computer Science, Georgia Institute of Technology, Atlanta, GA 30332, USA

Received 23 August 1983
Revised 22 May 1984

We consider directed graphs which have no short cycles. In particular, if n is the number of vertices in a graph which has no cycles of length less than $n-k$, for some constant $k<\frac{1}{3}n$, then we show that the graph has no more than 3^k cycles. In addition, we show that for $k\le\frac{1}{2}n$, there are graphs with exactly 3^k cycles. We thus are able to show that it is possible to bound the number of cycles possible in a graph which has no cycles of length less than $f(n)$ by a polynomial in n if and only if $f(n)\ge n-r\log(n)$ for some r.

1. Introduction

Undirected graphs with no short cycles (called graphs with large *girth*) have been the subject of much research over the years. However, digraphs with large girth are far less studied. It seems that no one has tried to determine how many cycles it is possible to have in a digraph with no short cycles. Our attention was directed toward this problem because of the following considerations: in some computer graphics applications, if the data are stored as a directed graph, certain transformations could be carried out by processing each cycle in the graph [2]. This is, of course, not practical in general, since a directed graph on n vertices can have as many as

$$\sum_{k=1}^{n}\binom{n}{k}(k-1)!$$

cycles. However, it is often possible to guarantee that the graphs under consideration have no short cycles. We are thus led to make the following definition, and to ask the following question:

Definition. Let f be any non-negative function on the positive integers. Then we define $C(f,n)$ to be the largest number of cycles possible in an n-vertex graph with no cycle of length less than $f(n)$.

*This research was supported by the National Science Foundation under Grant No. MCS 81-03608.

0166-218X/85/$3.30

Question. How big must $f(n)$ be for $C(f, n)$ to be bounded by a polynomial in n?

One of our results in this paper is in some respects a negative result, since we answer the above question by showing that $C(f, n)$ is polynomial only if $f(n) \geq n - r\log(n)$ for some constant r.

We also show that the converse implication holds; i.e., that $f(n) \geq n - r\log(n)$ implies that $C(f, n)$ is polynomial. This will be shown to follow from our main result; if $3k < n$, and if a graph has no cycle of length less than $n-k$, then the graph has no more than 3^k cycles. In addition, we prove that such a graph can have no more than 2^k cycles of length $n-k$.

In the next section, we give a lower bound for the number of cycles possible in a graph with no cycles of length less than $n-k$, for $2k \leq n$. In the following section, we give an upper bound for cycles of length exactly $n-k$, if $3k < n$. In Section 4, we build on the result of Section 3 to prove our main result. In the final section, we give remarks and conclusions.

Note that by *cycles* we mean *elementary cycles* (also called *simple cycles*); i.e., a cycle is a path from some vertex v back to v in which no vertex other than v is visited twice. All graphs mentioned will be taken to be directed graphs. Logarithms will be assumed to be base 2.

2. A lower bound

Theorem 2.1. *Let $2k \leq n$. Then there is an n-vertex graph with no cycle of length less than $n-k$, which has 3^k cycles, of which 2^k are of length $n-k$.*

Proof. Given vertices $v_1, v_2, \ldots, v_n$, insert the following edges:

$$
\begin{array}{ll}
v_i \to v_{i+1} & \text{for } 2k+1 \leq i \leq n-1, \\
v_i \to v_{k+i} & \text{for } 1 \leq i \leq k, \\
v_i \to v_{k+i+1}, v_i \to v_{i+1}, v_{k+i} \to v_{k+i+1}, v_{k+i} \to v_{i+1} & \text{for } 1 \leq i \leq k-1, \\
v_k \to v_{2k+1}, v_{2k} \to v_{2k+1}, v_n \to v_1, v_n \to v_{k+1} & \text{if } 2k < n, \\
v_k \to v_1, v_k \to v_{k+1}, v_{2k} \to v_1, v_{2k} \to v_{k+1} & \text{if } 2k = n.
\end{array}
$$

Some graphs which result from this construction are shown in Figs. 1 and 2.

It is easy to see that these graphs have no cycle of length less than $n-k$. To see that there are 2^k cycles of length $n-k$, consider the k 'columns' which have two vertices in them, and note that we can choose either the top or the bottom vertex for inclusion in any cycle of length $n-k$. To get longer cycles, we use the vertical edges. Thus at each of the k columns, we can either choose to 1) include the top vertex, 2) include the bottom vertex, or 3) include both; hence we can see that the graph has 3^k cycles. $\square$

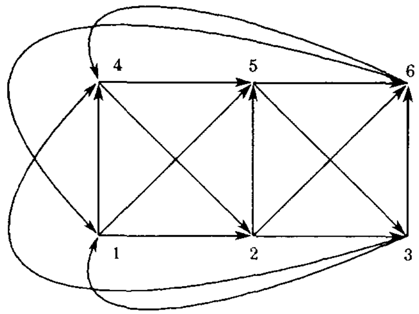

Fig. 1. $n=6$, $k=3$.

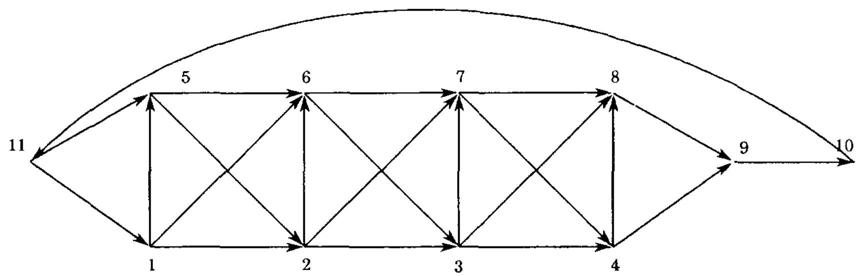

Fig. 2. $n=11$, $k=4$.

Corollary 2.2. *If $C(f, n)$ is bounded by a polynomial in n, then $f(n) \geq n - r\log(n)$ for some constant r.*

Proof. Assume that it is not true for any constant r that $f(n) \geq n - r\log(n)$. That is,

$$\forall r\, \exists n\colon \quad f(n) < n - r\log(n).$$

There are two cases.

Case 1: $\forall r\, \exists n\colon \quad (n \geq 2r\log(n) \wedge f(n) < n - r\log(n))$.

Let s be any integer, and let $r = 2s$. Recall that $\log(n) \geq \frac{1}{2}\lceil \log(n) \rceil$. By assumption, it follows that there exists an n such that

$$n \geq 2r\log(n) \geq r\lceil \log(n) \rceil = 2s\lceil \log(n) \rceil,$$

and

$$f(n) < n - r\log(n) \leq n - \tfrac{1}{2}r\lceil \log(n) \rceil = n - s\lceil \log(n) \rceil.$$

By Theorem 2.1, there is an n-vertex graph with no cycle shorter than $f(n)$ but which has $3^{s\lceil \log(n) \rceil} \geq n^s$ cycles. It follows easily that $C(f, n)$ is bounded by no polynomial in n.

Case 2: $\exists r\, \forall n\colon \quad (n \geq 2r\log(n) \Rightarrow f(n) \geq n - r\log(n))$.

Let m be the least integer such that $m \geq 2r\log(m)$. Note that $n > m$ implies that

$n \geq 2r\log(n)$. Thus for all $n \geq m$, $f(n) \geq n - r\log(n) \geq n - mr\log(n)$. By choice of m, we have that $n < m$ implies $n < 2r\log(n)$. If there is any n less than m, we also clearly have $m \geq 2$. Thus for all $n < m$, $f(n) \geq 0 \geq n(1 - \frac{1}{2}m) = n - m\frac{1}{2}n \geq n - mr\log(n)$. Thus for all n, $f(n) \geq n - mr\log(n)$. But this is contrary to our assumption that there is no constant r such that $f(n) \geq n - r\log(n)$. $\square$

The rest of this paper is devoted to showing that directed graphs with large girth must look very much like the examples given above if they have as many cycles as possible. In fact, the only way in which they may differ from those examples is that the 'columns' which have two vertices need not be adjacent. As an example, consider Fig. 3.

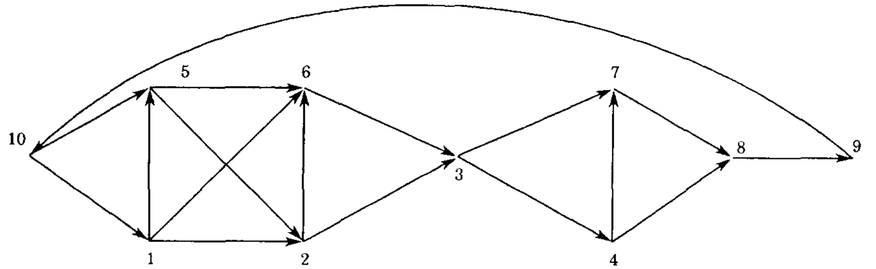

Fig. 3. A 10-vertex graph with girth $10-3$ having 3^3 cycles, of which 2^3 are of length 7, which differs from the graphs constructed in the proof of Theorem 1.

3. A preliminary result

We will denote cycles by lists of vertices. For instance, let $V = v_1, v_2, \ldots, v_{n-k}$ be a cycle of length $n-k$. (We may say that V is an *$(n-k)$-cycle*.) If $V' = v_i, v_{i+1}, \ldots, v_{n-k}, v_1, v_2, \ldots, v_{i-1}$, we say that V' is a *rotation* of V (or we say that we *rotate* V to get V'). Note that V and V' represent the same cycle; when counting cycles, we do not distinguish between rotations of a cycle. We shall also think of cycles as sets of vertices; for example, we may write $v_1 \in V$. If $v \notin V$, we say that v *is avoided by V*.

The next lemma gives some justification for considering cycles as sets of vertices, since it easily follows from Lemma 3.1 that, given a set of $n-k$ vertices, there is at most one $(n-k)$-cycle through them.

Lemma 3.1. *Let $V = v_1, v_2, \ldots, v_{n-k}$ and $W = w_1, w_2, \ldots, w_{n-k}$ be cycles in a graph in which no cycle has length less than $n-k$. If, for any j, $v_j = w_j$, then for all i, $1 \leq i \leq n-k$, we have that $v_i \in W$ implies that $v_i = w_i$.*

Proof. Let $v_j = w_j$. Rotate W and V so that $v_1 = w_1$. We will show, by induction on i, that for all positions i, either $v_i = w_i$, or $v_i \notin W$. (The claim is true for $i = 1$.) Assume the claim holds for $j < i$, but not for i. Thus we may assume that $v_i \neq w_i$, and $v_i = w_{i+l}$ for some $l>0$. But then $v_1, \dots, v_{i-1}, w_{i+l}, \dots, w_{n-k}$ is a cycle of length less than $n-k$, which is a contradiction. Thus the claim holds for all positions i. □

Taken informally, Lemma 3.1 says that two $(n-k)$-cycles can only differ in a few places. We now wish to generalize Lemma 3.1 in such a way that we can talk about a larger set of cycles differing in only a few places. To this end, we make the following

Definition. Let C be any set of $(n-k)$-cycles. An *arrangement* of C is a $|C| \times (n-k)$ matrix A such that for all i, $A_{i,1}, A_{i,2}, \dots, A_{i,n-k}$ is a cycle in C, and for every cycle V in C, $V = A_{i,1}, A_{i,2}, \dots, A_{i,n-k}$ for some i. (In that case, we will say that $V = A_i$.)

Lemma 3.2. *Let C be a set of $(n-k)$-cycles in a graph with no cycle of length less than $n-k$. If $3k<n$, then there exists an arrangement A of C such that for all i and j: $A_{i,m} = A_{j,n}$ implies that $m=n$. (I.e., a vertex can appear in at most one column of A.)*

Proof. By induction on $|C|$.

Basis: $|C|=2$. Let V and W be two $(n-k)$-cycles. Since $n-k > \frac{1}{2}n$, there is some vertex in $V \cap W$. Rotate V and W so that $v_1 = w_1$. Now by Lemma 3.1, we have that the lemma holds in this case.

Induction step: Let $|C| = r-1$. By the induction hypothesis, there is an arrangement A for C such that for all i and j : $A_{i,m} = A_{j,n}$ implies that $m=n$. We will show how to add a row to A to get an arrangement for $C \cup \{V\}$, where V is any $(n-k)$-cycle not in C. Since $3k<n$, A_1 and V have some vertex (say $A_{1,j}$) in common. Rotate V, if necessary, and add a row A_r which represents V, such that $A_{rj} = A_{1,j}$. By Lemma 3.1, for all i such that $A_{1,i}$ is in V, $A_{1,i} = A_{r,i}$. Now let A_s be any other cycle in A. Since $3k<n$, there is some vertex common to A_1, A_r, and A_s. Assume it is $A_{r,l}$. By the induction hypothesis applied to A, $A_{1,l} = A_{s,l}$. By Lemma 3.1 applied to A_1 and A_r, $A_{r,l} = A_{1,l}$. Thus $A_{r,l} = A_{s,l}$, and by Lemma 3.1 $A_{s,i} = A_{r,i}$ for all i such that $A_{s,i}$ is in V. Thus, we obtain an arrangement for $C \cup \{V\}$ satisfying the condition of the Lemma. By induction the result follows. □

Lemma 3.3. *Let A be an arrangement of $(n-k)$-cycles as in Lemma 3.2. Then if $P = \{ j \mid \exists n, m : A_{m,j} \neq A_{n,j} \}$, then $|P| \leq k$. (I.e., there are at most k columns which have more than one vertex.)*

Proof. There are k vertices which are not in A_1. Each of these vertices can appear in at most one column of A, by Lemma 3.2. For all of the other columns, only the vertex in A_1 can appear in that column. □

Theorem 3.4. *Let $3k<n$. If G is a graph with no cycle of length less than $n-k$, then G has no more than 2^k cycles of length $n-k$.*

Proof. By Lemmas 3.2 and 3.3, there is an arrangement A of the $(n-k)$-cycles in G such that at least $n-2k$ columns of A have only one vertex, and no vertex appears in more than one column. Let us assume without loss of generality that there is only one vertex v in the first column of A. It is clear that all $(n-k)$-cycles in G go through v, and that any $(n-k)$-cycle going through v must include one vertex from each column. Thus, if column j has c_j distinct vertices in it, there are clearly no more than B $(n-k)$-cycles, where

$$B=\prod_{j=1}^{n-k} c_j$$

How big can B be? Consider the k vertices not in A_1. Ignoring the rest of A for the moment, we can distribute those k vertices in any way we wish. Clearly, we can achieve this distribution by first placing the k vertices in k different columns, and then moving the vertices one by one, always moving vertices from columns with two vertices to columns with $h \geq 2$ vertices. Let B_i be the total number of cycles possible with the vertices arranged as they are after move i. $B_0=2^k$, since there are k columns with two vertices, and $n-2k$ columns with one vertex. Now consider the ith move, where we move a vertex from a column with two vertices to a column with h vertices. Then

$$B_{i+1}=\frac{h+1}{2h} B_i<B_i,$$

and $\max\{B_i : i \geq 0\}=B_0=2^k$. Thus, the maximum number of $(n-k)$-cycles occurs when each of k columns contains exactly two vertices. □

4. An upper bound

Now we will show how to adapt the proof of Theorem 3.4 to deal with cycles of arbitrary length (greater than or equal to $n-k$, if $3k<n$). We do this in five stages:

First, we assign a column number to each vertex.

Second, we show that for every edge $v \to w$, either

(1) v is in some column j, and w is in column $j+1 \pmod{n-k}$, (*Note:* here, and in the rest of the paper, we use the expression '$x \pmod{n-k}$' to mean 'the unique integer i in the set $\{1,2,\ldots,n-k\}$ such that $x \equiv i \pmod{n-k}$'. No confusion should result from this double use of the notation '$(\bmod\, n-k)$'.)

(2) v and w are in the same column, or

(3) v is in column j, and w is in column $j-r$, for some $r, 1 \leq r \leq k-1$, and the columns $j-r$, $j-r+1, \ldots, j$ each have at least two vertices in them. (We will call edges of this sort *Type 3* edges.)

Third, we show that there is a natural ordering which we can place on the vertices. This ordering has properties which prove useful in later stages.

Fourth, we show that we can maximize the number of cycles only if there are no

Type 3 edges.

Fifth, we show that we can maximize the number of cycles only if no column has more than two vertices. At that point, the proof of our main result will be complete.

Before we do any of that, however, let us prove a simple extension of lemma 3.1.

Lemma 4.1. *Suppose that $3k<n$ and that G is a graph which has n vertices and no cycles of length less than $n-k$. If V and W are any two cycles, then V and W can be rotated so that $V=v_1, v_2, \dots, v_r$, $W=w_1, w_2, \dots, w_s$, $v_1=w_1$, and for all i and j, if $v_i=w_h$, $v_j=w_l$, and $i<j$, then $h<l$. (I.e., vertices appear in the same order in V and W; hence for any given set of vertices, there is at most one cycle which goes through exactly the vertices in that set.)*

Proof. First note that since $r>\frac{1}{2}n$ and $s>\frac{1}{2}n$, there is some vertex in both V and W. Rotate V and W so that $v_1=w_1$. Now assume that $v_i=w_h$, $v_j=w_l$, $i<j$ and $h>l$. Note that

$v_1, \dots, v_{i-1}, w_h, \dots, w_s$ is a cycle of length $(i-1)+(s-h)+1=s+(i-h)\geq n-k$.

$w_1, \dots, w_{l-1}, v_j, \dots, v_r$ is a cycle of length $(l-1)+(r-j)+1=r+(l-j)\geq n-k$.

$w_l, \dots, w_h, v_{i+1}, \dots, v_{j-1}$ is a cycle of length $(h-l)+1+(j-1)-(i+1)+1=(j-l)+(h-i)\geq n-k$.

Adding, we get $r+s\geq 3n-3k>2n$. But $r+s\leq 2n$. $\square$

The results in this section deal with counting the number of cycles in a given graph G. Since those vertices of G which are not in any cycles can be deleted without affecting the number of cycles in the resulting graph, let us assume throughout this section that all of the vertices of G appear on cycles. Making this assumption has the effect of simplifying some of our definitions.

In section three, we proved Theorem 3.4 by showing how each vertex could be assigned a 'column number'. In discussions of cycles which may have length greater than $n-k$, we will still find it useful to think of arranging vertices into columns. Given a graph G which has n vertices and no cycle of length less than $n-k$, for some constant $k<\frac{1}{3}n$, let an *array* of G be an assignment of the vertices of G into $n-k$ columns labelled $1, 2, \dots, n-k$. Let $\mathrm{col}(v)$ be the column label of the column containing v, for each vertex in G.

If Z and Z' are both arrays of G (where $\mathrm{col}(v)$ gives the column of vertex v in array z, and $\mathrm{col}'(v)$ gives the column of vertex v in array Z') and there is some constant i such that for all vertices v, $\mathrm{col}(v)\equiv \mathrm{col}'(v)+i \pmod{n-k}$, then Z' is a *rotation* of Z.

Given a graph G in which the shortest cycle has length $n-k$ *and* $k<\frac{1}{3}n$, the following algorithm will create an array Z, which will be used throughout the rest of this section:

(1) Let A be an arrangement of the $(n-k)$-cycles in the form guaranteed by lemma 3.2. For every vertex v appearing in column j in A, set $\mathrm{col}(v)=j$.

(2) *phase* := 1

Let P = the set of cycles of G which contain vertices v such that col(v) is not yet defined.

while $P \neq \emptyset$

Let $X = x_1, x_2, \ldots, x_r$ be a cycle of shortest length in P.

(Let v be a vertex in X such that col(v) has not yet been defined. In the following, we will say that X *places* v. We shall also need to refer to the order in which vertices are placed. If we say that v is placed *in phase* q, we shall mean that v is placed during a pass through the **while** loop when the variable *phase* is equal to q. Phase 0 corresponds to step 1.) Note that more than half of the n vertices are in X, and more than half of the vertices are in the original arrangement A. Thus assume without loss of generality that x_1 is placed in an earlier phase

for $i := 1$ to $|X| - 1$

if x_{i+1} is not placed

then $\text{col}(x_{i+1}) := \text{col}(x_i) + 1 \pmod{n-k}$

end for

$\mathit{phase} := \mathit{phase} + 1$

$P := P - \{W \mid W$ is a cycle in G and W contains only vertices which have been placed$\}$

end while

Note that no vertex is assigned to more than one column. Thus there are at least $n - 2k > \frac{1}{2}(n-k)$ columns which have only one vertex, and hence there must be two adjacent columns which have only one vertex.

(3) Take the array which results after step 2 and rotate it so that columns 1 and $n-k$ each have only vertex. The resulting array is Z.

The next few lemmas show why the array Z is useful.

Lemma 4.2. *Every cycle contains at least one vertex which appears in a column by itself.*

Proof. A cycle avoids at most k vertices. There are no fewer than $n - 2k \geq k + 1$ vertices appearing in columns with only one vertex. □

Definition. Let v and w be any two vertices. A *direct path* from v to w is a path $x_1, x_2, \ldots, x_r$, where $r \leq n-k$, $x_1 = v$, $x_r = w$, and the following conditions hold for i in the range $1 \leq i \leq r-1$:

(1) $x_i \rightarrow x_{i+1}$ is an edge in G, and

(2) $\text{col}(x_{i+1}) \equiv \text{col}(x_i) + 1 \pmod{n-k}$.

Lemma 4.3. *Let w be any vertex, and let v be any vertex appearing in a column by itself. Then there is a direct path from v to w, and there is no shorter path from v to w.*

Proof. By induction on the phase q in which vertex w is placed. (Note that all vertices v which appear alone in a column are placed in phase 0.)

Basis: w is placed in phase 0. Verification of the basis is left to the reader.

Induction step: Let $W = w_1, w_2, \ldots, w_i, w_{i+1}, \ldots, w_{i+l}, w_{i+l+1}, \ldots, w_r$, where W places $w_{i+1}, \ldots, w_{i+l}$, vertex w_i is placed during an earlier phase and $w \in \{w_{i+1}, \ldots, w_{i+l}\}$. By Lemma 4.2, W contains some vertex w_c which appears in a column by itself. Note that w_c is placed in phase 0. We can assume without loss of generality that $c = 1$. By the induction hypothesis, for every vertex v which appears in a column by itself, there is a direct path from v to w_i, and there is no shorter path from v to w_i. By the algorithm for placing vertices, there is a direct path from w_i to w_{i+j} for $1 \le j \le l$. Thus there is a direct path from v to w_{i+j} for $1 \le j \le l$. (In fact, it follows that there is a direct path from v to every vertex in W.) It remains only to show that there is no shorter path from any such v to any of the w_{i+j}.

Case 1: *W includes a vertex from every column.*

In this case $v = w_h$ for some h, and $W = w_h, w_{h+1}, \ldots, w_i, w_{i+1}, \ldots, w_{i+l}, w_{i+l+1}, \ldots, w_{h-1}$. (If $h = 1$, interpret w_{h-1} to be w_r.) By the algorithm, since W places w_{i+j}, $1 \le j \le l$, there is no cycle shorter than W which contains w_{i+j}. Thus there is no path from w_h to w_{i+j} which is shorter than (length of $w_h, w_{h+1}, \ldots, w_i$) $+ j$, which is no shorter than (length of a direct path from w_h to w_i) $+ j$, which is equal to the length of a direct path from w_h to w_{i+j}.

Case 2: *There is some column m which is avoided by W.*

Let Z' be the rotation of Z in which $\mathrm{col}'(w_1) = 1$. Note that there is a direct path from w_1 to w_r. (It should be emphasized that we do *not* have that W contains a direct path from w_1 to w_r.) Since there is an edge from w_r to w_1, it follows that there must be a cycle of length $\mathrm{col}'(w_r)$. Since G has no cycle of length less than $n - k$, it follows that $\mathrm{col}'(w_r) = n - k$. Since there is some column m' which is avoided by W, it follows that there must be some h, $1 \le h \le r - 1$ such that $\mathrm{col}'(w_h) < m' < \mathrm{col}'(w_{h+1})$. Note that w_{h+1} is placed in an earlier phase, since if that were not the case, we would have $\mathrm{col}'(w_h) + 1 = \mathrm{col}'(w_{h+1})$. But since there is a direct path from w_1 to w_h, there is a path which is shorter than a direct path from w_1 to w_{h+1}, in contradiction to the induction hypothesis. □

Corollary 4.4. *There is no edge* $v \to w$ *such that* $\mathrm{col}(v) + 1 < \mathrm{col}(w)$.

Proof. Recall that column 1 has only one vertex. Existence of such an edge $v \to w$ would imply the existence of a path shorter than a direct path from column 1 to w. □

Corollary 4.5. *There is no edge* $v \to w$ *such that* $\mathrm{col}(w) = \mathrm{col}(v) - r$, $r \le n - k - 2$, *and for some* i, *where* $0 \le i \le r$, *column* $\mathrm{col}(v) - r + i$ *has just one vertex.*

Proof. If $i \ne 0$, existence of such an edge $v \to w$ would imply the existence of a path shorter than a direct path from column $\mathrm{col}(v) - r + i$ to w. If $i = 0$, existence of such an edge $v \to w$ would imply the existence of a cycle of length (length of a direct path from w to v) $+ 1 \le n - k - 1$. □

Corollary 4.6. *For any edge $v \leftarrow w$, either*

(1) $\mathrm{col}(v) = j$ *and* $\mathrm{col}(w) \equiv j+1 \pmod{n-k}$ ($v \to w$ *is a Type* 1 *edge*),

(2) $\mathrm{col}(v) = \mathrm{col}(w)$ ($v \to w$ *is a Type* 2 *edge*), *or*

(3) $\mathrm{col}(v) = j$ *and* $\mathrm{col}(w) = j-r$, *for some* r, $1 \le r \le k-1$, *and the columns* $j-r$, $j-r+1, \ldots, j$ *each have at least two vertices in them* ($v \to w$ *is a Type* 3 *edge*).

Proof. This follows from the previous two corollaries, and from the fact that there are at most k columns with two or more vertices. □

Lemma 4.7. *In any array in which all edges are of Type* 1, 2, *or* 3, *every possible cycle includes vertices from every column.*

Proof. First we show that every cycle must include the vertices in columns 1 and $n-k$. Assume that some cycle W avoids one or both of columns 1 and $n-k$. Let the lowest- and highest-numbered columns visited by W be l and h, respectively; thus there is a path in G leading from column h to column l, but avoiding one or both of columns 1 and $n-k$. By Corollary 4.6, every column from l to h has at least two vertices. Thus l and h must be in a block of consecutive columns, each of which has at least two vertices. How many vertices can appear in this block of columns? Since there are at most k columns with two or more vertices, there are at least $n-2k$ columns with exactly one vertex each. Thus there are at most $n-(n-2k) = 2k$ vertices which appear in columns with other vertices. That is, W is entirely contained in a block of consecutive columns which consists of at most $2k$ vertices. Since $2k < n-k$, this contradicts our choice of G as having no cycle of length less than $n-k$.

Now note that if all edges are of Type 1, 2, or 3, there is no way to move from left to right except by going one column at a time. The lemma follows. □

Definition. Let $\{x_1, x_2, \ldots, x_n\}$ be the vertex set of G. Define the binary relation $\lessdot$ on $\{x_1, x_2, \ldots, x_n\}$ in the following way: $x_i \lessdot x_j$ if and only if $i \ne j$ and

(1) $\mathrm{col}(x_i) = 1$, or

(2) $\mathrm{col}(x_i) \ne 1$, and there is a path from x_i to x_j which avoids column 1, or

(3) $\mathrm{col}(x_i) \ne 1$, there is no path from x_i to x_j which avoids column 1, and $\mathrm{col}(x_i) < \mathrm{col}(x_j)$, or

(4) $\mathrm{col}(x_i) \ne 1$, there is no path from x_i to x_j which avoids column 1, $\mathrm{col}(x_i) = \mathrm{col}(x_j)$, and $i < j$.

Lemma 4.8. $\lessdot$ *is a total ordering on the vertex set of G such that*

(1) *If $v \to w$ is an edge, then either*

(1.i) $v \lessdot w$, *or*

(1.ii) $\mathrm{col}(v) = n-k$ *and* $\mathrm{col}(w) = 1$.

(2) *If* $\mathrm{col}(v)=\mathrm{col}(w)+1$, *and all paths from* v *to* w *go through column* 1, *then* $w \lessdot v$.

(3) *If there is a path from* v *to* w *which avoids column* 1, *then* $v \lessdot w$.

Proof. By Lemma 4.7, if we delete the edge from the vertex in column $n-k$ to the vertex in column 1, the resulting graph is a directed acyclic graph G'. We claim that it is obvious that $\lessdot$ is an extension of the partial order corresponding to G' such that for any two distinct vertices v and w of G, exactly one of $v \lessdot w$ and $w \lessdot v$ holds. Also, if $v \to w$ is an edge, then either $\mathrm{col}(v)=n-k$ and $\mathrm{col}(w)=1$, or $v \to w$ is an edge of G' (in which case, $v \lessdot w$). Points (2) and (3) are now obvious. □

Lemma 4.9. *Assume that G has as many cycles as any digraph with n vertices and no cycle of length less than* $n-k$. *Then for every pair of vertices* v *and* w, *where* $v \lessdot w$ *and* $\mathrm{col}(v)+1 \geq \mathrm{col}(w)$, *there is an edge* $v \to w$.

Proof. Assume $v \lessdot w$, $\mathrm{col}(v)+1 \geq \mathrm{col}(w)$, and there is no edge $v \to w$. It is easy to show that adding the edge $v \to w$ will increase the number of cycles. Thus the lemma holds if we can show that we have not altered the fact that no cycle is of length less than $n-k$. Note that if we add the edge $v \to w$, either

(1) $\mathrm{col}(v)+1=\mathrm{col}(w)$ (in which case $v \to w$ is a Type 1 edge),

(2) $\mathrm{col}(v)=\mathrm{col}(w)$ (in which case $v \to w$ is a Type 2 edge), or

(3) $\mathrm{col}(v)>\mathrm{col}(w)$. By the definition of $\lessdot$, if $v \lessdot w$ and $\mathrm{col}(v)>\mathrm{col}(w)$, then there must be a path from v to w which avoids column 1. Each right-to-left edge on the path from v to w must be a Type 3 edge. It follows that each of the columns $\mathrm{col}(w), \mathrm{col}(w)+1, \ldots, \mathrm{col}(v)$ has at least two vertices in it
(in which case $v \to w$ is a Type 3 edge).

Now it follows from Lemma 4.7 that every cycle must contain a vertex from every column. Thus every cycle must have length no less than $n-k$. □

Lemma 4.10. *The number of cycles in G is maximum only if there is no edge of Type* 3.

Proof. Assume that G has a maximum number of cycles, and that there is a Type 3 edge $v \to w$, where $\mathrm{col}(v)=j$, and there is no Type 3 edge from any column greater than j. We will show how to modify G to get a graph with more cycles.

By Corollary 4.6, column j has $r \geq 2$ vertices $v_1, v_2, \ldots, v_r$ in it, where $v_1 \lessdot v_2 \lessdot \cdots \lessdot v_r$. Thus $j \notin \{1, n-k\}$, since those columns each contain only one vertex. Note that, since there are no Type 3 edges from any column to the right of column j, all paths from any vertex u to any vertex x, where $\mathrm{col}(u)>j$ and $\mathrm{col}(x) \leq j$, must go through column 1. Thus, by Lemma 4.8, if $\mathrm{col}(u)=j+1$, we have $v_i \lessdot u$ for $1 \leq i \leq r$. Since G has as many cycles as possible, we have by Lemma 4.9 that there is an edge $v_i \to u$ for $1 \le i \le r$. That is, all possible edges exist from column j to column $j+1$.

Let D be the number of simple paths from column $j+1$ to column $n-k$, and let E_i be the number of simple paths from column 1 to v_i. It follows that the total number of cycles in G is

$$D \sum_{i=1}^{r} E_i.$$

Let l be the largest integer for which there is a Type 3 edge $v_l \to w$. Note that w was placed by some cycle W, where, by Lemma 4.7, W contains a path from w through columns $j-1$ and j to column $n-k$. Thus there is some w' on that path, where $\mathrm{col}(w') = j-1$; similarly, that path contains some v_h in column j. Since there is a path from v_l to w' which avoids column 1, it follows by lemma 4.8 that $v_l \ll w'$; it also follows that $v_l \ll v_h$. Thus we have that $l \ne r$ and $v_l \ll v_{l+1}$. Also note that $w \ll v_{l+1}$, since if it were true that $v_{l+1} \ll w$, it would follow from Lemma 4.9 that there is a Type 3 edge $v_{l+1} \to w$, but there is no such edge by the choice of l.

Let $\{w_1, w_2, \ldots, w_s\}$ be the set of all vertices of G which come between v_l and v_{l+1} in the ordering $\ll$. That is, let $v_l \ll w_1 \ll w_2 \ll \cdots \ll w_s \ll v_{l+1}$. Now by Lemma 4.9, there is an edge $v_l \to w_i$ for $1 \le i \le s$. The discussion in the preceding paragraph shows that at least one of the w_i is in column $j-1$.

We claim that G' has more cycles than G, where G' is obtained from G by

(1) reversing each edge $v_l \to w_i$ in G where $\mathrm{col}(w_i) = j-1$, and

(2) deleting each edge $v_l \to w_i$ in G where $\mathrm{col}(w_i) \ne j-1$.

First note that any cycle in G' which is not in G must contain an edge $w_i \to v_l$ for some i. However, it is clear that any path from v_l to w_i in G' must include column 1. It follows easily that G' has no cycle of length less than $n-k$.

Define D' and E_i' analogously to D and E_i above, so that the number of cycles in G' is

$$D' \sum_{i=1}^{r} E_i'.$$

We claim that it is obvious that $D' = D$, that $E_i' = E_i$ for $1 \le i < l$, and that $E_l' > E_l$. We also claim that for $i > l$, $E_i' = E_i$. To see this, first note that any path in G from column 1 to v_i which avoids v_l is also a path in G', as is any path from column 1 to v_i which includes v_l but avoids all vertices in the set $\{w_1, \ldots, w_s\}$. Now let $y_1, \ldots, y_a, v_l, w_b, \ldots, w_c, v_d, \ldots, v_i$ be any simple path in G from column 1 to v_i which includes v_l, where $\{w_b, \ldots, w_c\} \subseteq \{w_1, \ldots, w_s\}$, and $\{v_d, \ldots, v_i\} \subseteq \{v_{l+1}, \ldots, v_i\}$. Note that $\mathrm{col}(w_c) = j-1$, and thus there is an edge $w_c \to v_l$ in G'. Also note that $y_a \ll w_b$ and $\mathrm{col}(y_a) = j-1$ and $\mathrm{col}(w_b) \le j-1$; thus by Lemma 4.9, there is an edge $y_a \to w_b$ in G, and hence also in G'. But then $y_1, \ldots, y_a, w_b, \ldots, w_c, v_l v_d, \ldots, v_i$ is a path in G'. That completes the proof of our claim that $E_i' = E_i$ for $i > l$.

Thus,

$$D' \sum_{i=1}^{r} E_i' > D \sum_{i=1}^{r} E_i.$$

That is, G' has more cycles than G, a contradiction. $\square$

Theorem 4.11. *Let $3k<n$, and let H have no cycle shorter than $n-k$. Then G has no more than 3^k cycles.*

Proof. By Lemma 4.10, we can assume that G can be arranged in an array Z such that, for any edge $v \to w$, either

(1) $\mathrm{col}(v)+1 \equiv \mathrm{col}(w) (\mathrm{mod}\ n-k)$, or

(2) $\mathrm{col}(v)=\mathrm{col}(w)$.

We claim that the following are obvious:

(1) Every cycle must go through each column in order from left to right.

(2) Every cycle includes a non-empty subset of the vertices in each column.

(3) If two cycles are different, then there is some column in which they contain different vertices. (This holds because of Lemma 4.1.)

It follows that, if column i contains c_i vertices, then there are at most F cycles, where

$$F=\prod_{i=1}^{n-k}(2^{c_i}-1).$$

As in the proof of Theorem 3.4, we see that we can arrive at any assignement of vertices to columns by a series of moves, where we start with k columns with two vertices, and with each move we take a vertex from a column with two vertices and place it in a column with $h \geq 2$ vertices. Let F_i denote the greatest number of cycles possible when the vertices are arranged as after move i. F_0 is thus 3^k, and

$$F_{i+1}=\frac{(2^{h+1}-1)}{3(2^h-1)}F_i.$$

Thus $F_{i+1}<F_i$, and $\max\{F_i : i \geq 0\}=F_0=3^k$.

Thus the maximum number of cycles is achieved when each of k columns contains exactly two vertices. $\square$

Corollary 4.12. *$C(f,n)$ is bounded by a polynomial in n if and only if $f(n) \geq n-r\log(n)$ for some constant r.*

Proof. The only if direction was proved as Corollary 2.2. Thus assume that $f(n) \geq n-r\log(n)$. Let $g(n)$ denote $n-f(n)$. Thus $g(n) \leq r\log(n)$. Let M be such that $3r\log(n)<n$ for all $n>M$.

Let G be any graph with n vertices and no cycle of length less than $f(n)$. If $n>M$, then by Theorem 4.11 G has no more than $3^{g(n)} \leq 3^{\lfloor r\log(n)\rfloor} \leq n^{r\log(3)}$ cycles. Thus $C(f,n) \leq n^{\lceil r\log(3)\rceil}$ for almost every n, and hence there is some constant K such that $C(f,n) \leq n^{\lceil r\log(3)\rceil}+K$. $\square$

Note that we have essentially characterized the extremal graphs (i.e., those graphs having 3^k cycles and having no cycle of length less than $n-k$). It follows easily from the proofs presented above that the extremal graphs are those graphs presented in the proof of Theorem 2.1, without the restriction that the columns with two vertices be adjacent. (The details are left to the reader.)

5. Remarks and conclusions

We have as many questions as we have answers. In particular, we have:

Question 1. How many cycles is it possible to have, if $3k \geq n$?

Question 2. Given a random graph G, what is the probability that it has no 'short' cycles, or that it has 'not very many' cycles?

In answer to Question 1, we conjecture that for $2 \leq \frac{1}{3}n \leq k \leq \frac{1}{2}n$ it is impossible to have more than 3^k cycles. We plan to investigate Question 2 at some time in the future, but have little to say about it now. It is known, for instance, that if a random graph on n vertices has $o(n)$ edges, the probability that it has any cycles approaches zero [1, Theorem 3.a]. Thus it may be of some interest to know, for instance, if a random graph on n vertices has $f(n)$ edges, how small $f(n)$ must be in order to make it likely that a graph has a polynomial number of cycles.

A final note: Carsten Thomassen has also recently investigated digraphs with large girth; in [3] he gives a structural characterization of digraphs with girth at least two-thirds n. Using that characterization, he is able to give alternative proofs of the main results of this paper. He goes on to show that for $2 \leq \frac{1}{3}n \leq k \leq \frac{1}{2}n$ it is impossible to have more than $2^k(n-k)$-cycles, and for $k > \frac{1}{2}n > 3$, it is not possible to have $2^k(n-k)$-cycles.

Acknowledgements

I would like to thank Bruce Naylor for bringing this problem to my attention. Rich DeMillo suggested Question 2, and encouraged me to publish this result. Thanks are also due to Richard Duke, Kim King, Barbara Smith-Thomas, Jerry Spinrad, Marc Graham, and the referee for their helpful comments, which greatly improved the readability of this paper. Finally, I wish to thank K.B. Reid, Chris Godsil, Frank Harary, Brian Alspach, and especially Carsten Thomassen for assisting me in my efforts to discover what was already known about this question.

References

[1] P. Erdös and A. Rényi, On the evolution of random graphs, Magyar Tud. Akad. Mat. Kut. Int Közl. 5 (1960) 17–61. Also in: Joel Spencer, ed., Paul Erdös: The Art of Counting (The MIT Press Cambridge, MA, 1973) 574–617.
[2] Bruce Naylor, Personal communication.
[3] Carsten Thomassen, The 2-linkage problem for acyclic digraphs, Preprint (1983).